RESEARCH DEPARTMENT

Graphical solution of cubic and quartic equations

MATHEMATICAL REPORT No. F-025

1965/44

THE BRITISH BROADCASTING CORPORATION
ENGINEERING DIVISION

ADDENDUM

RESEARCH DEPARTMENT - BRITISH BROADCASTING CORPORATION

Mathematical Report No. F-025 (1965/44)

GRAPHICAL SOLUTION OF CUBIC AND QUARTIC EQUATIONS

This Report is the property of the British Broadcasting Corporation and may not be reproduced in any form without the written permission of the Corporation.

ADDENDUM

RESEARCH DEPARTMENT - BRITISH BROADCASTING CORPORATION

Mathematical Report No. F-025 (1965/44)

GRAPHICAL SOLUTION OF CUBIC AND QUARTIC EQUATIONS

1. INTRODUCTION

The original Report F-025 gave not only graphical methods of obtaining reasonable approximations to the roots of cubic and quartic equations, but also discussed iterative methods of improving the accuracy. The procedure suggested by Porter and Mack[1], however, appears to be superior, and this was not appreciated at the time of writing the original report.

The procedure of Porter and Mack is here explained with numerical examples, in the hope that an engineer will thus be able confidently to solve a quartic equation to the degree of accuracy usually required in practice. Butler[2] has given a clear account of more refined methods of achieving high accuracy. The Addendum thus has no originality; its object is merely to make other people's work easier.

The method can be extended to equations of higher degree, but attention here is confined to quartics, since the biggest 'hurdle' is usually to realise that a straightforward procedure exists at all for solving equations of degree above 2.

The above-mentioned procedure in effect permits any equation of degree up to 5 to be solved, since an equation of odd degree has at least one real root which can be found relatively easily (as indicated in Section 4 of the original report) and 'factored out'.

2. EXAMPLES OF THE METHOD OF PORTER AND MACK

Consider the quartic equation

$$x^4 + c_1x^3 + c_2x^2 + c_3x + c_4 = 0 \tag{1}$$

This can be factorized into

$$(x^2 + ax + b)\ (x^2 + lx + m) = 0 \tag{2}$$

or

$$x^4 + (a + l)x^3 + (b + al + m)x^2 + (bl + am)x + bm = 0 \tag{3}$$

Equating coefficients of Equation (3) with those of the original quartic (1) we have the following relations

$$a + l = c_1 \tag{4}$$

$$b + al + m = c_2 \tag{5}$$

$$bl + am = c_3 \tag{6}$$

$$bm = c_4 \tag{7}$$

The method of Porter and Mack[1] consists in choosing a likely value b_r for b, and then solving Equations (4), (6) and (7) for the corresponding values of a, l, m giving

$$m_r = c_4/b_r \tag{8}$$

$$a_r = \frac{c_1 b_r - c_3}{b_r - m_r} \tag{9}$$

$$l_r = c_1 - a_r \tag{10}$$

where a_r, l_r, m_r are the corresponding values of a, l, m when $b = b_r$ respectively.

Equation (2) is then written as

$$g_r = b_r + a_r l_r + m_r - c_2 \tag{11}$$

We seek to choose b_r so that g_r is zero. In other words, we choose b_r, calculate a_r, l_r and m_r as if this value were correct, and then the value of g_r tells us how close we are to the correct value of b.

Consider first the case when c_4 is positive. Test to see whether $c_4 = (c_3/c_1)^2$. This is the only 'special' case, as Butler[2] has pointed out; we then have

$$b = m = \sqrt{c_4} \tag{12}$$

while a and l are found from

$$a + l = c_1 \tag{13}$$

$$b + al + m = c_2 \tag{14}$$

so that

$$al = c_2 - 2\sqrt{c_4} \tag{15}$$

and

$$a - l = (c_1^2 - 4al)^{\frac{1}{2}}$$

$$= \{c_1^2 - 4c_2 + 8\sqrt{c_4}\}^{\frac{1}{2}} \tag{16}$$

$$\therefore \quad a = \tfrac{1}{2}\{c_1 + [c_1^2 - 4c_2 + 8\sqrt{c_4}]^{\frac{1}{2}}\}$$

$$l = c_1 - a \tag{17}$$

Example 1

$$x^4 + 5x^3 + 8x^2 + 5x + 1 = 0 \tag{18}$$

c_4 is positive and equal to $(c_3/c_1)^2$

$$\therefore \quad b = m = \sqrt{c_4} = 1 \tag{19}$$

$$a = \tfrac{1}{2}\{5 + (25 - 32 + 8)^{\frac{1}{2}}\} = 3 \tag{20}$$

$$l = 5 - 3 = 2 \tag{21}$$

giving

$$\left.\begin{aligned}&(x^2 + 3x + 1)\,(x^2 + 2x + 1)\\ &= x^4 + 5x^3 + 8x^2 + 5x + 1\end{aligned}\right\} \tag{22}$$

In general, however, $c_4 \neq c_3^{\,2}/c_1^{\,2}$ and there is a positive value of b between 0 and $\sqrt{c_4}$.

Take as first trial value of b, $\tfrac{1}{2}\sqrt{c_4}$. Calculate m, a, l and g from (8), (9) and (10).

If g is positive, take next value of b as $\tfrac{3}{4}\sqrt{c_4}$.

If g is negative, take next value of b as $\tfrac{1}{4}\sqrt{c_4}$.

Calculate m, a, l and g again from (8), (9) and (10). If both values of g are positive, take next value of b as $0{\cdot}9\sqrt{c_4}$.

If both values of g are negative, take next b as $0{\cdot}1\sqrt{c_4}$.

If the 2 values of g are of opposite sign, then the next b to try is calculated by linear interpolation, i.e.

$$b_3 = \frac{b_1 g_2 - b_2 g_1}{g_2 - g_1} \tag{23}$$

where b_1, b_2 and b_3 are the successive b's, g_1 and g_2 being the g's associated with b_1, b_2.

Then calculate m, a, l and g from (8), (9) and (10) with $b = b_3$.

If the g's are still remaining either all positive or all negative, the b's to be tried are respectively $0{\cdot}99\sqrt{c_4}$, $0{\cdot}999\sqrt{c_4}$ etc. and $0{\cdot}01\sqrt{c_4}$, $0{\cdot}001\sqrt{c_4}$ etc. Otherwise, linear interpolation is used to find the successive b's. The process in continued until g is small enough for the accuracy required.

Example 2

$$x^4 + 2x^3 + 10x^2 + 5x + 1 = 0 \tag{23}$$

c_4 is positive, and not equal to c_3^2/c_1^2 so we try first

$$b = \tfrac{1}{2}\sqrt{c_4} = 0{\cdot}5$$

The working is best set out as in Table 1.

TABLE 1

$c_1 = 2, \quad c_2 = 10, \quad c_3 = 5, \quad c_4 = 1$

b_r	m_r	$b_r - m_r$	$c_1 b_r - c_3$	a_r	l_r	g_r
0•5	2	-1•5	-4	2•67	-0•67	-9•29
0•25	4	-3•75	-4•5	1•2	0•8	-4•79
0•1	10	-9•9	-4•8	0•485	1•515	0•835
0•12	8•333	-8•213	-4•76	0•5796	1•420	-0•724
0•111	9•009009	-8•898009	-4•778	0•536974	1•463026	-0•094384
0•1097	9•115770	-9•006070	-4•7806	0•530820	1•469180	0•005340
0•109770	9•109957	-9•000187	-4•78046	0•531151	1•468849	-0•000092

This gives for the 2 quadratic factors of the left hand side of (23)

$$(x^2 + 0{\cdot}531151x + 0{\cdot}109770)\,(x^2 + 1{\cdot}468849x + 9{\cdot}109957)$$

$$= x^4 + 2{\cdot}000000x^3 + 9{\cdot}999908x^2 + 4{\cdot}999998x + 1{\cdot}000000 \qquad (24)$$

It should be noted that there is no need to work to many significant figures until g becomes small.

Example 3

(Illustrating the case when the g's remain positive for longer than is usual).

$$x^4 + 5x^3 + 8{\cdot}01x^2 + 5{\cdot}03x + 1{\cdot}01 = 0 \qquad (25)$$

c_4 is positive, and nearly but not exactly equal to c_3^2/c_1^2.

Try first, $b = \tfrac{1}{2}\sqrt{c_4} = 0{\cdot}5$. The working is as in Table 2.

TABLE 2

$c_1 = 5, \quad c_2 = 8{\cdot}01, \quad c_3 = 5{\cdot}03, \quad c_4 = 1{\cdot}01$

b_r	m_r	$b_r - m_r$	$c_1 b_r - c_3$	a_r	l_r	g_r
0·5	2	-1·5	-2·5	1·7	3·3	0·1
0·75	1·35	-0·6	-1·28	2·13	2·87	0·203
0·91	1·1099	-0·1999	-0·48	2·4012	2·5988	0·2501
0·995	1·015	-0·020	-0·055	2·75	2·25	0·1875
1·0040	1·0060	-0·0020	-0·0100	5·00	0	-6·00

The last entry in Table 2 indicates that g_r is varying violently in this region because $b_r \approx m_r$. Linear interpolation may not therefore be satisfactory as a means of obtaining the next b_r, but the correct b must be between 0·995 and 1·004. As an alternative to linear interpolation, therefore, the next b_r can always be taken as the average of any two b_r associated with g_r of opposite sign, and in this way the interval within which b must lie can be progressively reduced until eventually linear interpolation is adequate for the degree of accuracy required. In this particular case the required quadratic factors are exactly $(x^2 + 3x + 1)\,(x^2 + 2x + 1{\cdot}01)$.

Next consider the case when c_4 is negative. We first determine

$$H = (c_1^2/4) + (c_3^2/4c_4) - c_2 \tag{26}$$

If $H>0$, there is a value of b between 0 and $\sqrt{|c_4|}$. Take as first value of b, $\tfrac{1}{2}\sqrt{|c_4|}$.

Calculate m, a, l and g from (8), (9) and (10) as usual.

If g is positive, take next value of b as $\tfrac{1}{4}\sqrt{|c_4|}$.

If g is negative, take next value of b as $\tfrac{3}{4}\sqrt{|c_4|}$.

Recalculate m, a, l and g.

If both values of g are positive, take next value of b as $0{\cdot}1\sqrt{|c_4|}$.

If both values are negative, take next b as $0{\cdot}9\sqrt{|c_4|}$.

If the two values are of opposite sign, the next b is obtained by linear interpolation.

Recalculate m, a, l and g.

If the g's are still remaining all positive, the b's to be tried are $0{\cdot}01\sqrt{|c_4|}$, $0{\cdot}001\sqrt{|c_4|}$ etc. Otherwise, linear interpolation can be used, or the alternative averaging procedure mentioned in Example 3.

Example 4

$$x^4 + 9x^3 + 2x^2 + 5x - 2 = 0 \tag{27}$$

c_4 is negative. $H = 15{\cdot}125>0$

So we try, $b = \frac{1}{2}\sqrt{|c_4|} = 0{\cdot}7$

$c_1 = 9,\ c_2 = 2,\ c_3 = 5,\ c_4 = -2$

TABLE 3

b_r	m_r	$b_r - m_r$	$c_1b_r - c_3$	a_r	l_r	g_r
0•7	-2•86	3•56	1•3	0•37	8•63	-0•97
1•06	-1•887	2•947	4•54	1•541	7•459	8•667
0•74	-2•703	3•443 .	1•66	0•482	8•518	0•143
0•735	-2•7211	3•4561	1•615	0•46729	8•5327	0•00115
0•73496	-2•72124	3•45620	1•61464	0•467172	8•53283	0•000019

which gives for the two quadratic factors required

$$(x^2 + 0{\cdot}467172x + 0{\cdot}73496)(x^2 + 8{\cdot}53283x - 2{\cdot}72124)$$

$$= x^4 + 9{\cdot}000002x^3 + 2{\cdot}000019x^2 + 5{\cdot}000002x - 2{\cdot}000003 \qquad (28)$$

If H is negative, there is a value of b between $-\sqrt{|c_4|}$ and 0. Take as first value of b, $-\frac{1}{2}\sqrt{|c_4|}$.

Calculate m, a, l and g using (8), (9) and (10) as usual.

If g is positive, take next value of b as $-\frac{3}{4}\sqrt{|c_4|}$.

If g is negative, take next value of b as $-\frac{1}{4}\sqrt{|c_4|}$.

Recalculate m, a, l and g.

If both values of g are positive, take next value of b as $-0{\cdot}9\sqrt{|c_4|}$.

If both values of g are negative, take next b as $-0{\cdot}1\sqrt{|c_4|}$.

If the 2 values are of opposite sign, the next b is obtained by linear interpolation.

Recalculate m, a, l and g.

If the g's are still remaining all negative, the b's to be tried are $-0{\cdot}01\sqrt{|c_4|}$, $-0{\cdot}001\sqrt{|c_4|}$ etc. respectively. Otherwise linear interpolation is used, or the averaging procedure mentioned in Example 3.

Example 5

$$x^4 - 5x^3 + 7x^2 + 9x - 4 = 0 \qquad (29)$$

c_4 is negative. $H = \frac{25}{4} - \frac{81}{16} - 7 < 0$

So we try $b = -\frac{1}{2}\sqrt{|c_4|} = -1$

$c_1 = -5, \quad c_2 = 7, \quad c_3 = 9, \quad c_4 = -4$

TABLE 4

b_r	m_r	$b_r - m_r$	$c_1b_r - c_3$	a_r	l_r	g_r
-1	4	-5	-4	0•8	-5•8	-8•64
-0•5	8	-8•5	-6•5	0•76	-5•76	-3•88
-0•2	20	-20•2	-8	0•396	-5•396	10•6632
-0•42	9•524	-9•944	-6•9	0•694	-5•694	-1•8476
-0•388	10•30927	-10•69727	-7•06	0•659981	-5•659981	-0•81421
-0•3628	11•025358	-11•388158	-7•186	0•631006	-5•631006	0•109359
-0•36578	10•935535	-11•301315	-7•1711	0•634536	-5•634536	-0•005561
-0•365636	10•939841	-11•305477	-7•17182	0•634366	-5•634366	-0•000045

This gives for the 2 quadratic factors

$$(x^2 + 0{\cdot}634366x - 0{\cdot}365636)\,(x^2 - 5{\cdot}634366x + 10{\cdot}939841) \tag{30}$$

$$= x^4 - 5{\cdot}000000x^3 + 6{\cdot}999955x^2 + 8{\cdot}999990x - 3{\cdot}9999997$$

If $H = 0$, we can take $b = -\sqrt{|c_4|}$, $m = +\sqrt{|c_4|}$, a and l are then found from Equations (4) and (5) which give

$$\left.\begin{aligned} a &= \tfrac{1}{2}\{c_1 + (c_1^2 - 4c_2)^{\frac{1}{2}}\} \\ l &= c_1 - a \end{aligned}\right\} \tag{31}$$

Example 6

$$x^4 + 6x^3 + 7x^2 + 4x - 2 = 0 \tag{32}$$

$c_4 < 0 \qquad H = 0$

$\therefore \qquad b = -\sqrt{2} = -1{\cdot}414214$

$m = +1{\cdot}414214$

$a = 4{\cdot}414214$

$l = 6 - 4{\cdot}414214 = 1{\cdot}585786$

giving

$$(x^2 + 4{\cdot}414214x - 1{\cdot}414214)\,(x^2 + 1{\cdot}585786x + 1{\cdot}414214)$$

$$= x^4 + 6{\cdot}000000x^3 + 6{\cdot}999999x^2 + 4{\cdot}000002x - 2{\cdot}000001 \qquad (33)$$

3. REFERENCES

1. A. PORTER *and* C. MACK. 1949. New methods for the numerical solution of algebraic equations, *Phil.Mag.* 1949. Series 7, **40**, pp. 578 - 585.

2. R. BUTLER. 1962. The rapid factorization to high accuracy of quartic expressions, *Am.math.Mon.*, 1962. **69**, 2, pp. 138 - 141.

JWH/SMW
29/4/66

Printed by BBC Research Department, Kingswood Warren, Tadworth, Surrey

RESEARCH DEPARTMENT

GRAPHICAL SOLUTION OF CUBIC AND QUARTIC EQUATIONS

Mathematical Report No. F-025

(1965/44)

D. Maurice

J.W. Head, M.A., F.Inst.P., A.M.I.E.E., F.I.M.A.

for Head of Research Department

This Report is the property of the British Broadcasting Corporation and may not be reproduced in any form without the written permission of the Corporation.

Mathematical Report No. F-025

GRAPHICAL SOLUTION OF CUBIC AND QUARTIC EQUATIONS

Section	Title	Page
	SUMMARY	1
1.	INTRODUCTION	1
2.	CUBIC EQUATIONS	3
3.	QUARTIC EQUATIONS	6
4.	IMPROVING THE ACCURACY OF A GRAPHICAL FIRST APPROXIMATION BY ITERATIVE PROCEDURES	11
5.	CONCLUSIONS	12
6.	REFERENCES	13

November 1965 Mathematical Report No. F-025

(1965/44)

GRAPHICAL SOLUTION OF CUBIC AND QUARTIC EQUATIONS

SUMMARY

Given any cubic or quartic equation with real coefficients, it can be transformed by a change of variable into a standard form, and thus solved by means of a double-sided scale (Fig. 1) in the case of a cubic, or by means of the nomogram of Fig. 3 (or Fig. 4) in the case of a quartic. Alternative procedures for certain cases where high accuracy with this nomogram is difficult to obtain are also discussed. Thus the information here given enables cubic and quartic equations to be solved as easily (to an accuracy of say two significant figures) as quadratic equations are solved without this information. Once a good approximation to the roots is known, accuracy can be improved by iterative processes if necessary; this is briefly discussed. The normal routine required for solving a cubic or quartic to graphical accuracy is specified in the Conclusions (Section 5).

1. INTRODUCTION

Various well-known methods are available for solving algebraic equations with real coefficients, including special ones which are available only for cubic or quartic equations. Most of these involve either iterative procedures (which can be tedious if no approximate prior information about the roots is available) or some transformation of the given equation into a special form, containing parameters which have to be determined by means of an auxiliary calculation. It is reasonably easy to obtain approximate values of real roots by drawing a graph of the left hand side of the equation, but in the case of an equation of even degree all of whose roots are complex, there is no obvious method of locating complex roots approximately except the now well-known 'root-locus' method which again is straight-forward in principle, but usually requires a considerable amount of calculation before an adequate approximation to the roots being sought becomes available.

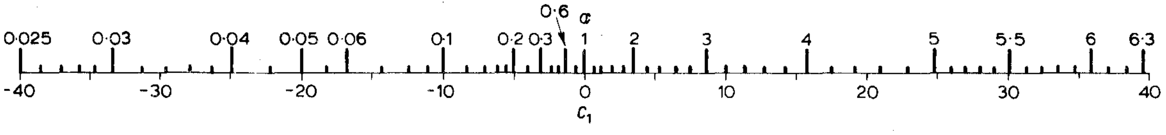

Fig. 1 - Double-sided scale for finding the Positive Real Root of Equation (5)

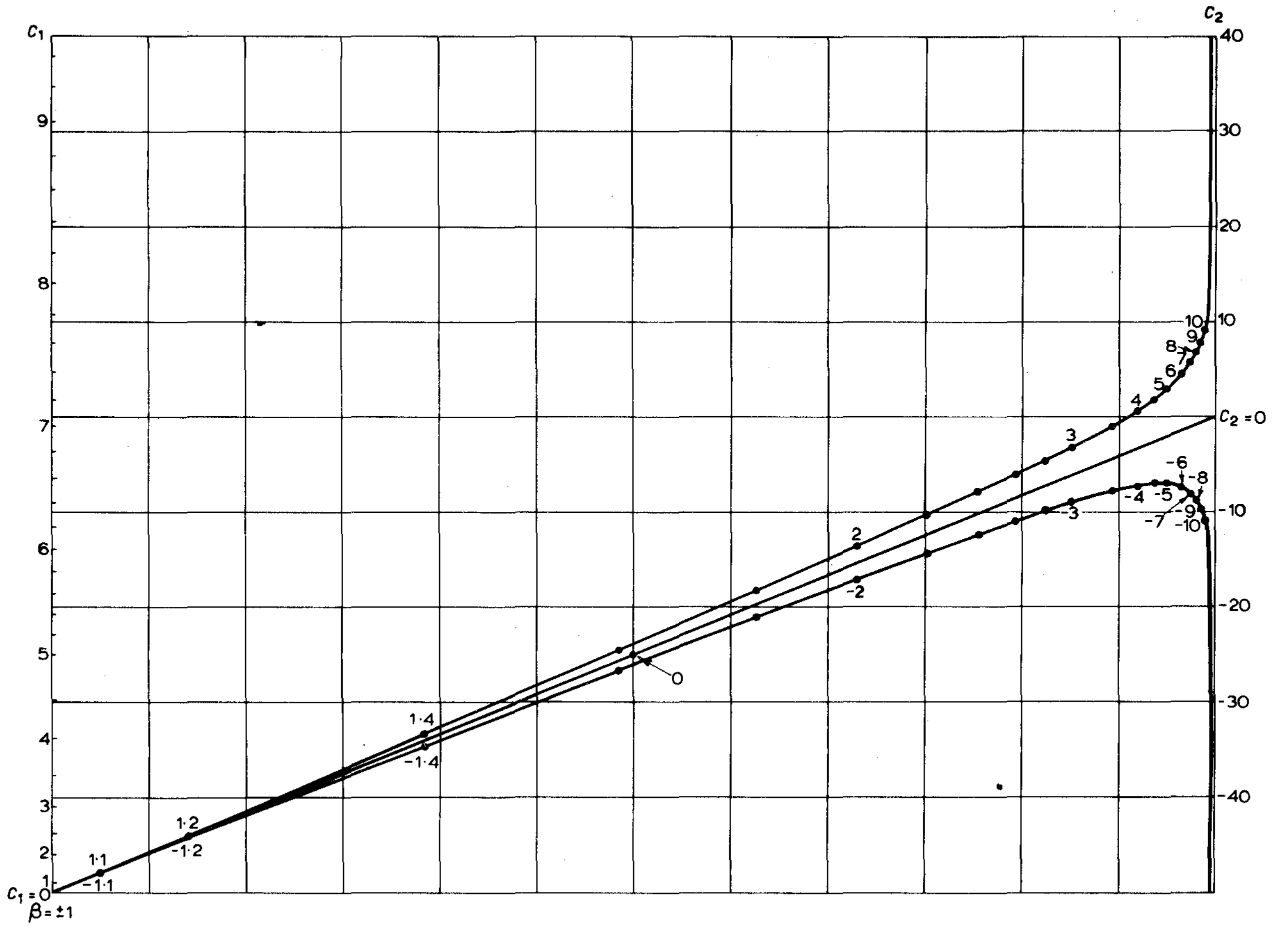

Fig. 2 - Nomogram for finding the number β in expression (18) associated with the factors of the left hand side of equation (16)

The general method of attack here considered is to use a two-sided scale (Fig. 1) to obtain the real root of a cubic, and a nomogram (Fig. 2, 3 or 4) to obtain the number-term of a real quadratic factor of a quartic. This is far from being the first time that nomograms have been used for solving algebraic equations, since d'Ocagne[1,2,3] gave techniques as long ago as 1902 for finding the real roots of equations of degrees up to 7, and Mack and Tomlin[4] constructed a nomogram for the solution of quartics in 1945.

Any algebraic equation in x can be 'normalized' by changes of variable so that there is no term in x^{n-1} if the equation is of degree n, and so that the term independent of x is ± 1. This normalization must be carried out before the techniques to be described are used. If the equation so normalized is cubic, the relation between the positive real root and the normalized coefficient of x is one that can be determined and expressed graphically as in Fig. 1 once for all as is explained in Section 2. If the equation so normalized is quartic, the relation between the coefficient of x^2, the coefficient of x and the number term in the quadratic factor being sought can be expressed in the form of a nomogram (Figs. 2, 3 or 4) as is explained in Section 3. Cases where there is at first sight some difficulty in using the nomogram satisfactorily are also discussed in Section 3. When a good approximation to a root of an equation is known, accuracy can be improved by means of iterative

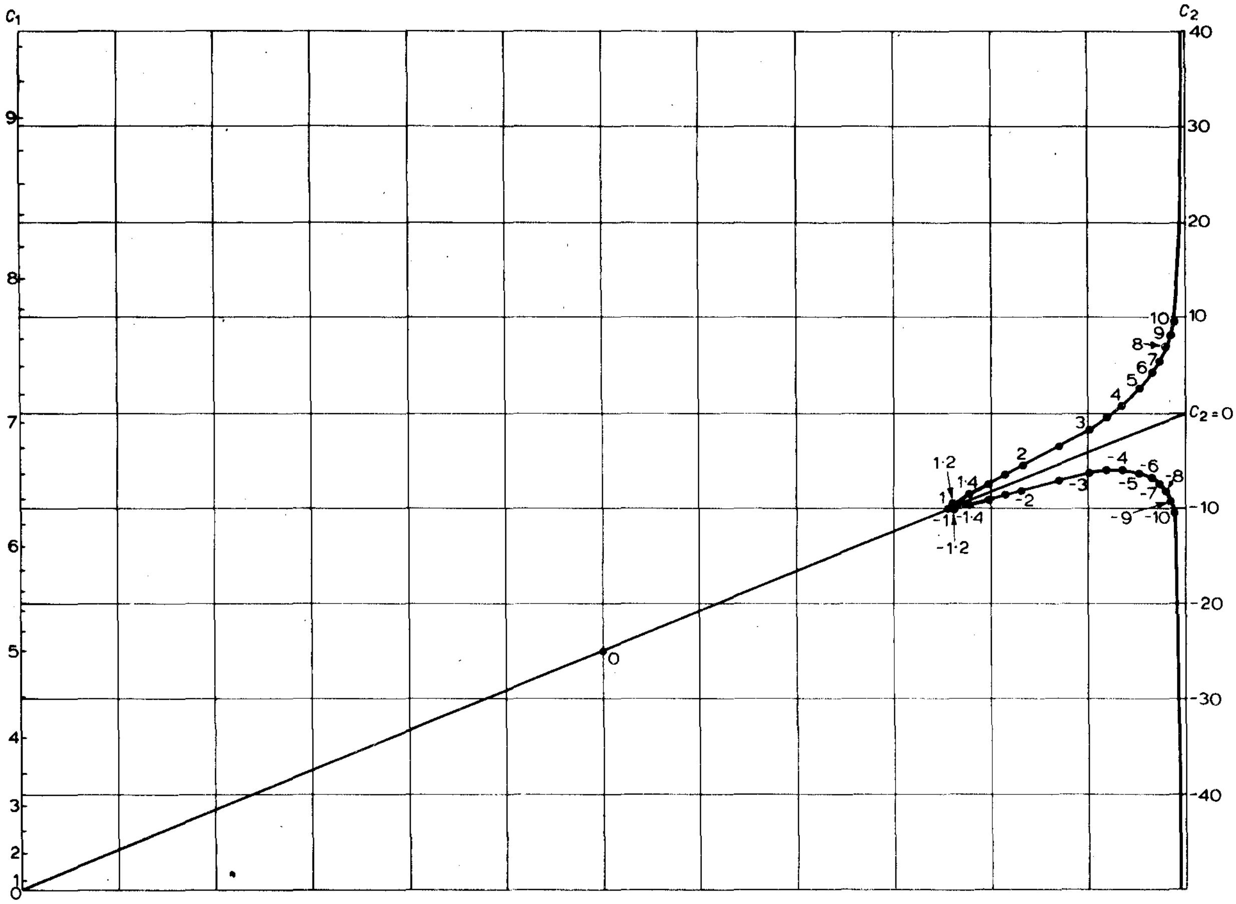

Fig. 3 - Nomogram analogous to Fig. 2 for the case when b_o in equation (14) is negative

procedures; this is briefly discussed in Section 4 and more fully in References 4, 5 and 6. Section 5 gives conclusions, in the form of a specification of the routine to be followed for solving a cubic or quartic.

2. CUBIC EQUATIONS

Suppose that the given equation is

$$x^3 + a_2 x^2 + a_1 x + a_o = 0 \tag{1}$$

then by writing

$$y = x + \frac{1}{3}a_2 \tag{2}$$

Equation (1) can be reduced to

$$y^3 + b_1 y + b_o = 0 \tag{3}$$

where

$$b_1 = a_1 - \frac{1}{3}a_2^2; \quad b_o = a_o - \frac{1}{3}a_1 a_2 + \frac{2}{27}a_2^3 \tag{4}$$

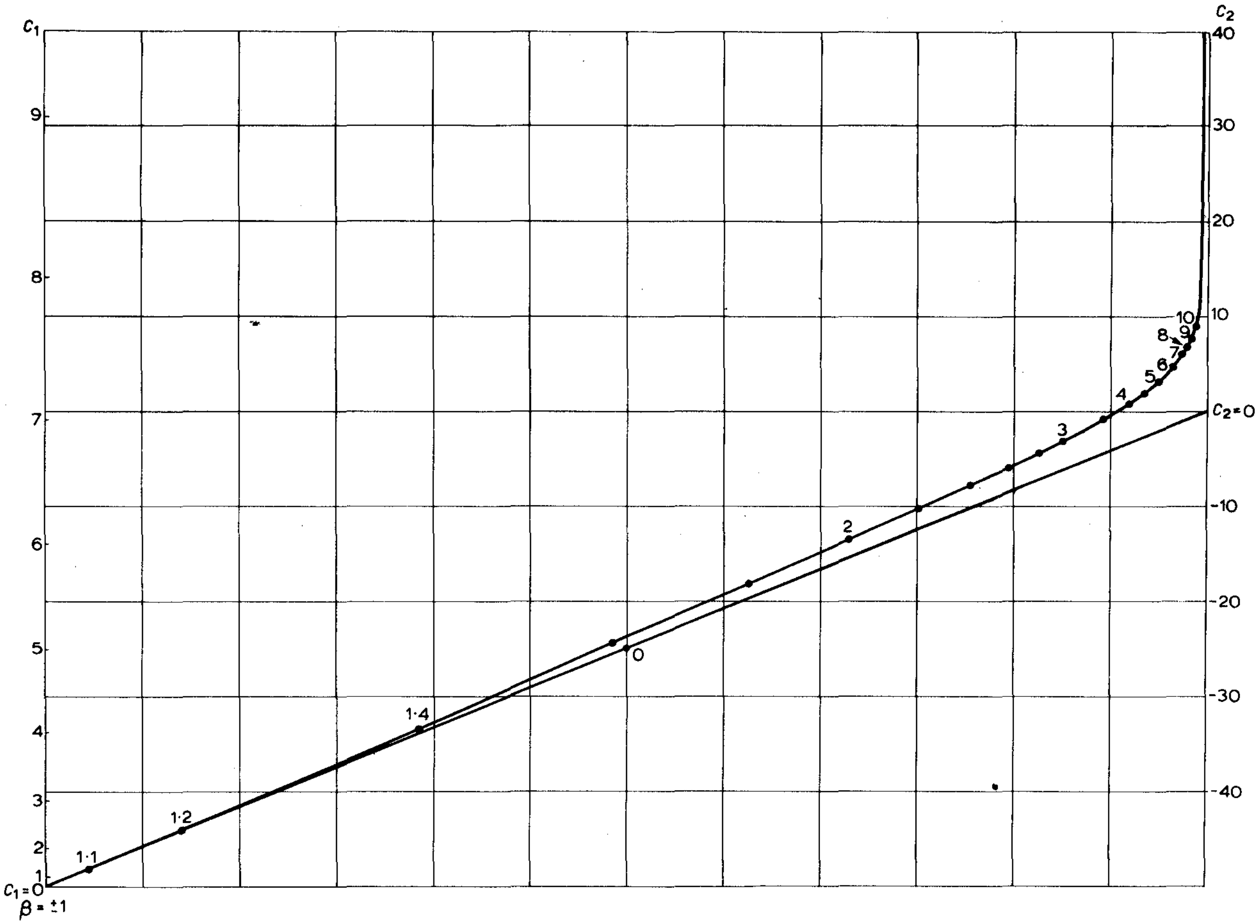

Fig. 4 - Nomogram identical with Fig. 2 except for the suppression of the negative-β branch

and finally y can be changed to $-b_0^{1/3}z$ so that the equation reduces to the normalized form

$$z^3 - c_1 z - 1 = 0 \qquad (5)$$

where

$$c_1 = -b_1/b_0^{2/3} \qquad (6)$$

If $b_0 \approx 0$, $x \approx -a_2/3$ is the required approximation to a real root of (1). The remaining factor of the left hand side of (1) can then be taken as

$$x^2 + \frac{2}{3}a_2 x + \frac{3a_0}{a_2}$$

Equation (5) necessarily has a positive root say α; the left-hand side is therefore expressible in the form

$$(z - \alpha)\ (z^2 + \alpha z + \frac{1}{\alpha}) \qquad (7)$$

and

$$c_1 = \alpha^2 - \frac{1}{\alpha} \tag{8}$$

Now it is easy to construct as in Fig. 1 a two-sided scale in which there is a linear c_1-scale on one side, and a non-linear scale of the corresponding values of α on the other side. From this double scale, the value of α corresponding to a given value of c_1 can be read, and the factors (7) of the left-hand side of the normalized equation (5) deduced immediately. It only remains to express factors in terms of the original variable x; to do this we must substitute

$$z = -\left(x + \frac{1}{3}a_2\right)/b_o^{1/3} \tag{9}$$

into (7). The factors of the left hand side of the original equation (1) thus becomes respectively

$$x + \frac{1}{3}a_2 + b_o^{1/3}\alpha = x + p, \text{ say} \tag{10}$$

and

$$x^2 + x\left(\frac{2}{3}a_2 - \alpha b_o^{1/3}\right) + \left[\frac{1}{9}a_2^2 - \frac{1}{3}a_2\alpha b_o^{1/3} + (b_o^{2/3}/\alpha)\right]$$

$$= x^2 + qx + r, \text{ say} \tag{11}$$

Examples:

In this and later sets of examples, cases with exact roots are chosen and the exact roots are given, so that the accuracy obtainable by the graphical technique is easily assessed.

(i) $x^3 + 9x^2 + 20x + 12 = 0$

or $(x + 1)(x + 2)(x + 6) = 0$

(ii) $x^3 + 7x^2 + 14x + 8 = 0$

or $(x + 1)(x + 2)(x + 4) = 0$

(iii) $x^3 + 6x^2 + 11x + 6 = 0$

or $(x + 1)(x + 2)(x + 3) = 0$

(iv) $x^3 + 3x^2 + 4{\cdot}25x + 3 = 0$

or $(x + 1{\cdot}5)(x^2 + 1{\cdot}5x + 2) = 0$

Example No.	a_2	a_1	a_o	b_1	b_o	c_1	α	p	q	r
(i)	9	20	12	-7	6	2·12	1·56	5·84	3·16	2·61
(ii)	7	14	8	-2·333	0·740	2·86	1·9	4·05	2·95	1·865
(iii)	6	11	6	-1	0	-	-	2	4	3
(iv)	3	4·25	3	1·25	0·75	-1·51	0·6	1·545	1·455	1·83

3. QUARTIC EQUATIONS

Suppose that the given quartic equation is

$$x^4 + a_3x^3 + a_2x^2 + a_1x + a_o = 0 \tag{12}$$

then by writing

$$y = x + \frac{1}{4}a_3 \tag{13}$$

it can be reduced to

$$y^4 + b_2y^2 + b_1y + b_o = 0 \tag{14}$$

where

$$\left.\begin{aligned} b_2 &= a_2 - 0{\cdot}375a_3^2 \\ b_1 &= a_1 - 0{\cdot}5a_2a_3 + 0{\cdot}125a_3^3 \\ b_o &= a_o - 0{\cdot}25a_1a_3 + 0{\cdot}0625a_2a_3^2 - 0{\cdot}01171875a_3^4 \end{aligned}\right\} \tag{15}$$

Consider first the case in which b_o is positive, and put $y = (b_o^{1/4})z$, and (14) reduces to

$$z^4 + c_2z^2 + c_1z + 1 = 0 \tag{16}$$

where

$$c_2 = b_2/b_o^{\frac{1}{2}}; \quad c_1 = b_1/b_o^{3/4} \tag{17}$$

We seek to factorize (16) - the factors will be of the form

$$(z^2 + \alpha z + \beta)\ (z^2 - \alpha z + \frac{1}{\beta}) \tag{18}$$

so that the z^3 term and the term independent of z are correct; in order that the z^2 and z terms of (16) may be also correct, we must have

$$\beta + \frac{1}{\beta} - \alpha^2 = c_2 \tag{19}$$

$$-\ \alpha(\beta - \frac{1}{\beta}) = c_1 \tag{20}$$

Eliminating α from (19) and (20) we have

$$\left[\beta + \frac{1}{\beta}\right] - \frac{c_1^2}{\left[\beta - \frac{1}{\beta}\right]^2} = c_2 \tag{21}$$

Now Equation (21) is of the form

$$f_1f_3 + f_2g_3 + h_3 = 0 \tag{22}$$

where

$$\left.\begin{array}{l} f_1 = c_1^2; \quad f_2 = c_2; \quad f_3 = 1/[\beta-(1/\beta)]^2 \\ g_3 = 1; \quad h_3 = -\ [\beta + (1/\beta)] \end{array}\right\} \qquad (23)$$

so that f_1 depends only on c_1, f_2 depends only on c_2, and f_3, g_3, h_3 depend only on β. Equation (22) is a standard form, which means that the relation can be expressed by means of a nomogram (see reference 7) - it is the condition that the points

$$\mathrm{P}\ (-d_o,\ \mu_1 f_1),\ \mathrm{Q}(d_o,\ \mu_2 f_2) \text{ and } \mathrm{R}(X_3,\ Y_3)$$

shall be in a straight line, where d_o, μ_1 and μ_2 are scale factors at our choice, and

$$X_3 = d_o \frac{\mu_1 g_3 - \mu_2 f_3}{\mu_1 g_3 + \mu_2 f_3}\ ; \quad Y_3 = -\frac{\mu_1 \mu_2 h_3}{\mu_1 g_3 + \mu_2 f_3} \qquad (24)$$

As c_1 varies, the point P just mentioned moves along a line parallel to the Y-axis, and the lowest point of the scale occurs for $c_1 = 0$, since $f_1 = c_1^2$. On the other hand, as c_2 varies, the point Q also moves parallel to the Y-axis, but the point corresponding to $c_2 = 0$ should be near the middle of this scale since f_2 is proportional to c_2 and may therefore have either sign. There are many possible choices of d_o, μ_1 and μ_2; it is convenient to have d_o (half the distance between the point representing $c_1 = 0$ and the point representing $c_2 = 0$) a round number, and to decide in advance the approximate overall size of the nomogram. Fig. 2 was therefore constructed originally to have an overall size of 240 mm by 180 mm, and with the zero of the c_2 - scale 100 mm above the lower edge of the paper, so that

$$\left.\begin{array}{l} d_o = ½\surd(240^2 + 100^2) = 130 \text{ mm}; \\ \mu_1 = \mu_2 = 2 \text{ mm per unit of corresponding } f \end{array}\right\} \qquad (25)$$

The X-axis is then the oblique line joining $c_1 = 0$ (bottom left-hand corner) to $c_2 = 0$ (about half way up the right hand side of the diagram) and the Y - axis is vertical. The origin O is midway between the vertical c_1 - scale on the extreme left and the vertical (and linear) c_2 - scale on the extreme right. The β-scale, defined by equations (24), is curved, and has two branches, according to whether β is positive or negative. The β-scale has been calibrated with a value of β (numerically greater than unity). Once a value of β has been found, Equation (20) gives the associated value of α and completes the factorization (18) to graphical accuracy.

If c_1 is small, it may not be easy to determine accurately the position of the intersection with the β-curve of the line joining the points $(-d_o,\ \mu_1 f_1)$ calibrated by the value of c_1 and the point $(d_o,\ \mu_2 f_2)$ calibrated by the value of c_2. When this happens, or if the index line meets the β-curve beyond the last calibration $\beta = \pm 10$, a starting approximation can be obtained by neglecting c_1, in which case (16) becomes quadratic in z^2 so that

$$\beta + \frac{1}{\beta} \approx c_2 \qquad (26)$$

Having found β from (26) a first approximation to α can be obtained from (20) and an improved value of β is then obtainable from (19) and an improved (and probably adequate) value of α from (20).

Equation (26) is unsatisfactory if $|c_2| < 2$. Instead, replace β initially by 1 in. (19) if $0 < c_2 < 2$ or by -1 if $0 > c_2 > -2$; this gives a first approximation to α, and (20) then gives an improved value of β which in turn can be substituted into (19) to give an improved value of α and so on. Alternatively, γ_o can be replaced by c_2 as an approximate root of (36) below.

If $c_1 = 0$ and c_2 is greater than 2, the quartic (16) has four purely imaginary roots; we can therefore expect (16) to have four complex roots when c_1 is sufficiently small and $c_2 > 2$. In Fig. 2 the index line (joining the point marked c_2 to the point marked c_2) in such a case intersects the positive-β branch of the β-curve in one point and does not meet the negative-β branch at all. The factor $(z^2 + \alpha z + \beta)$ thus derived will be associated with a complex root-pair of (16); the remaining factor $z^2 - \alpha z + (1/\beta)$ may be associated with two real roots of the same sign, or may be associated with a complex root-pair. If however c_2 is sufficiently negative, the possibility arises that the index line may meet the negative-β branch of the β-curve in two places as well as the positive-β branch in one. This is the case when (16) has four real roots, two positive and two negative, which can be combined in three different pairs as real quadratic factors of the left-hand side of (16).

As before, it is necessary to express the factors found in terms of the original variable x; this can be done by means of the substitution

$$z = (x + \frac{1}{4}a_3)/b_o{}^{1/4} \tag{27}$$

and the real quadratic factors of the left hand side of (12) thus become

$$\begin{aligned} & x^2 + (\tfrac{1}{2}a_3 + \alpha b_o{}^{1/4})x + (\frac{1}{16}a_3{}^2 + \frac{1}{4}a_3\alpha b_o{}^{1/4} + \beta b_o{}^{\frac{1}{2}}) \\ & = x^2 + p_1 x + q_1, \text{ say} \end{aligned} \tag{28}$$

and

$$\begin{aligned} & x^2 + (\tfrac{1}{2}a_3 - \alpha b_o{}^{1/4})x + (\frac{1}{16}\, a_3{}^2 - \frac{1}{4}a_3\alpha b_o{}^{1/4} + [b_o{}^{\frac{1}{2}}/\beta]) \\ & = x^2 + p_2 x + q_2, \text{ say} \end{aligned} \tag{29}$$

Examples $[b_o > \text{o}]$

(v) $\quad x^4 + 4x^3 + 9x^2 + 8x + 5 = 0$

$\quad$ or $(x^2 + x + 1)\ (x^2 + 3x + 5) = 0$

(vi) $\quad x^4 + 3{\cdot}2x^3 + 4{\cdot}71x^2 + 3{\cdot}04x + 0{\cdot}8 = 0$

$\quad$ or $(x^2 + 1{\cdot}1x + 0{\cdot}4)\ (x^2 + 2{\cdot}1x + 2) = 0$

(vii) $\quad x^4 + 8x^3 + 22{\cdot}75x^2 + 29{\cdot}5x + 17{\cdot}5 = 0$

or $(x + 2\cdot5)\ (x + 3\cdot5)\ (x^2 + 2x + 2) = 0$

(viii) $x^4 + 10x^3 + 35x^2 + 50x + 24 = 0$

or $(x + 1)\ (x + 2)\ (x + 3)\ (x + 4) = 0$

Example No.	b_2	b_1	b_o	c_2	c_1	β	α	p_1	q_1	p_2	q_2
(v)	3	-2	3	1·732	-0·877	1·75	0·744	2·98	5·01	1·02	1·01
(vi)	0·87	-0·4	0·1536	2·220	-1·630	2·1	1·00	0·97	0·33	2·23	1·96
(vii)	-1·25	2·5	1·5	-1·021	1·845	1·54	-2·072	1·71	1·30	6·29	9·38
(viii)	-2·5	0	0·5625	-3.1/3	0	-3	0	5	4	5	6

If b_o is negative, the relation between y and z is $y = (-b_o)^{1/4}z$, and (14) reduces to

$$z^4 + c_2z^2 + c_1z - 1 = 0 \tag{30}$$

where

$$c_2 = b_2/(-b_o)^{\frac{1}{2}}; \quad c_1 = b_1/(-b_o)^{3/4} \tag{31}$$

and the quadratic factors are of the form

$$(z^2 + \alpha z + \beta)\ (z^2 - \alpha z - \frac{1}{\beta}) \tag{32}$$

We then have

$$\beta - \frac{1}{\beta} - \alpha^2 = c_2 \tag{33}$$

$$-\alpha(\beta + \frac{1}{\beta}) = c_1 \tag{34}$$

so that the relation on which the nomogram for this case is based (obtained by eliminating α between (33) and (34)) is

$$(\beta - \frac{1}{\beta}) - \frac{c_1^{\,2}}{\left[\beta + \frac{1}{\beta}\right]^2} = c_2 \tag{35}$$

This is as before of the form (22) but

$$f_1 = c_1^{\,2}; \quad f_2 = c_2; \quad f_3 = 1/[\beta + (1/\beta)]^2 \tag{36}$$

$$g_3 = 1 \quad h_3 = (1/\beta) - \beta \ .$$

and with the definitions in Equations (36) instead of Equations (23), Equations (24) and (25) still apply to define the β-curve for the new nomogram (Fig. 3); the c_1 and c_2 scales are unchanged. In Fig. 3, as in Fig. 2, β is specified as numerically greater than 1; $\beta = \pm 1$ on the X-axis, and β is positive above this axis and negative below it. If there are three intersections of the index line with the β-curve, the quartic (30) has four real roots. If the index line meets the β-curve beyond $\beta = \pm 10$ the approximate value of β can be obtained from (33) with $\alpha = 0$.

Having determined β from Fig. 3 and deduced α from (34), we again have to express the factors found in terms of the original variable x; this requires the substitution

$$z = (x + \frac{1}{4}a_3)/(-b_o)^{1/4} \qquad (37)$$

and the real quadratic factor of the left hand side of (12) thus become

$$x^2 + (\tfrac{1}{2}a_3 + \alpha[-b_o]^{1/4})x + \left[\frac{1}{16}a_3{}^2 + \frac{1}{4}a_3\alpha(-b_o)^{1/4} + \beta(-b_o)^{\frac{1}{2}}\right] \qquad (38)$$

$$= x^2 + P_1x + Q_1 \text{ say}$$

and

$$x^2 + (\tfrac{1}{2}a_3 - \alpha[-b_o]^{1/4})\ x + \left[\frac{1}{16}a_3{}^2 - \frac{1}{4}a_3\alpha(-b_o)^{1/4} - (-b_o{}^{\frac{1}{2}})/\beta\right] \qquad (39)$$

$$= x^2 + P_2x + Q_2, \text{ say}$$

Examples $b_o < o$

(ix) $x^4 + 3{\cdot}2x^3 + 4{\cdot}41x^2 + 2{\cdot}41x + 0{\cdot}2 = 0$

or $(x^2 + 1{\cdot}1x + 0{\cdot}1)\ (x^2 + 2{\cdot}1x + 2) = 0$

(x) $x^4 + 8x^3 + 19x^2 + 22x + 10 = 0$

or $(x^2 + 6x + 5)\ (x^2 + 2x + 2) = 0$

(xi) $x^4 - 25x^2 - 60x - 36 = 0$

or $(x + 1)\ (x + 2)\ (x + 3)\ (x - 6) = 0$

(xii) $x^4 + 8x^3 + 27x^2 + 70x + 50 = 0$

or $(x + 1)\ (x + 5)\ (x^2 + 2x + 10) = 0$

Example No.	b_2	b_1	b_o	c_2	c_1	β	α	P_1	Q_1	P_2	Q_2
(ix)	0·57	-0·55	-0·1344	1·555	-2·478	2·0	0·991	2·2	1·85	1·0	-0·024
(x)	-5	10	-6	-2·041	2·608	-2·0	1·043	5·63	2·37	2·37	1·96
(xi)	-25	-60	-36	-4·167	-4·082	1·0	2·041	5	6	-5	-6
(xii)	3	26	-30	0·054	2·178	1·4	-1·03	1·60	6·8	6·40	4·91

4. IMPROVING THE ACCURACY OF A GRAPHICAL FIRST APPROXIMATION BY ITERATIVE PROCEDURES

The main difficulty in seeking complex roots of algebraic equations is to obtain a good starting approximation, and this we have succeeded in doing for cubics and quartics by the procedure of Sections 2 and 3. If we want greater accuracy than is available by graphical means, two alternative iterative methods of improving accuracy are briefly discussed below.

For a cubic equation, it is naturally the real root which should be found to a sufficient degree of accuracy in this way, and this root should then be factorized out, leaving only a residual quadratic to be solved in the usual way. If the real root has smaller modulus then either of the remaining roots, the processes to be described will be quite straightforward. If it has larger modulus then the remaining roots, it may be necessary to change z into $1/t$ in Equation (5), and actually solve the t-equation instead of the z-equation.

Suppose that $z \approx \alpha$ is the approximation already obtained graphically and that the accurate value of z is $\alpha + \delta$. Then we can write (5) in the form

$$(z - \alpha - \delta)\left[z^2 + (\alpha + \delta)\, z + \frac{1}{\alpha + \delta}\right] = 0 \tag{40}$$

and choose δ so that the coefficient of z in the product is correct, that is,

$$c_1 = (\alpha + \delta)^2 - \frac{1}{\alpha + \delta} \approx \alpha^2 - (1/\alpha) + \delta[2\alpha + (1/\alpha^2)] \tag{41}$$

If α is a good approximation, $c_1 \approx \alpha^2 - (1/\alpha)$ and thus δ is small; the process can be repeated with $(\alpha + \delta)$ instead of α if necessary. This procedure is effectively an adaptation of Bairstow's method.[5] An alternative process originally due to Lin[6] is even simpler in principle; we divide the left hand side of (5) by $z(z-\alpha)$; the remainder is

$$(\alpha^2 - c_1)\, z + 1 \tag{42}$$

and the next approximation to z is $1/(c_1 - \alpha^2) = \alpha'$ say. If the process is now repeated and a further approximation α'' thus obtained, then a very close approximation is given by Aitken's formula

$$z = \frac{\alpha'^2 - \alpha\,\alpha''}{2\alpha' - \alpha - \alpha''} = \alpha + \frac{\Delta^2{}_1}{2\Delta_1 - \Delta_2} = A \tag{43}$$

where

$$\Delta_1 = \alpha' - \alpha \qquad \Delta_2 = \alpha'' - \alpha \tag{44}$$

If necessary, the whole process can be repeated with A as the starting approximation instead of α.

When we are similarly trying to improve the approximation to a quadratic factor $z^2 + \alpha z + \beta$ of (16) or (30) the Lin procedure can be applied, not to (16) or (30) themselves, but to (21) or (35). In (21) put

$$\beta + (1/\beta) = \gamma \tag{45}$$

and (21) becomes

$$\gamma(\gamma^2 - 4) - c_2(\gamma^2 - 4) - c_1^2 = 0$$

or

$$\gamma^3 - c_2\gamma^2 - 4\gamma + (4c_2 - c_1^2) = 0 \tag{46}$$

We have an approximation β_0 to β from Fig. 2 or 4, and from this we can derive $\gamma_0 = \beta_0 + (1/\beta_0)$ which is approximately a root of (46). We can therefore divide the left hand side (46) by $\gamma(\gamma - \gamma_0)$ and find a better approximation analogous to A in (43) and thus obtain γ as accurately as we need. The corresponding value of β is obtained from the quadratic equation (45). In (35) put

$$\beta - (1/\beta) = \gamma \tag{47}$$

and (35) becomes

$$\gamma(\gamma^2 + 4) - c_2(\gamma^2 + 4) - c_1^2 = 0$$

or

$$\gamma^3 - c_2\gamma^2 + 4\gamma - (4c_2 + c_1^2) = 0 \tag{48}$$

and if β_0 is an approximation to β obtained from Fig. 3, $\gamma_0 = \beta_0 - (1/\beta_0)$ is an approximation to a root of (48), so the left hand side can be divided by $\gamma(\gamma - \gamma_0)$ and a better approximation analogous to A in (43) can be found as before. Equation (47) gives the corresponding value of β.

5. CONCLUSIONS

The following are the normal routine steps required to solve any cubic or quartic equation 'to graphical accuracy'.

A. Cubic Equations

(i) Convert the given equation (1) into the normalized form (5), using equations (4) and (6).

(ii) From Fig. 1 given c_1 from equations (6) and (4), read off α, a positive real root of (5).

(iii) Substitute into (10), that is find p in order to obtain the linear factor of the left hand side of (1), and into (11), that is, finding q and r in order to obtain the complementary quadratic factor, and check against (1) by multiplying out. The roots are then $x = -p$ and $x = -q \pm (q^2 - 4r)^{\frac{1}{2}}$.

B. Quartic Equations

(i) Convert the given equation (12) to the normalized form (14) by using equations (15) and again to the form (16) using equations (17) if $b_0>0$ or (30) using equations (31) if $b_0<0$.

(ii) Use the nomogram (Fig. 4 if $b_0>0$, Fig. 3 if $b_0<0$) to determine a value of β, the number term of the quadratic factor (18) or (32) being sought. (In Fig. 4 the negative-β part of Fig. 2 has been suppressed; this merely tells us whether the quartic has four real roots or not, and by factorizing the quadratics (18) obtained with positive β, this information is equally well available.)

(iii) Use Equation (20) or (34) to determine α, the coefficient of z in the quadratic factor (18) or (32) being sought.

(iv) Re-convert to the original variable x by substitution into (28) and (29) ($b_0>0$) or into (38) and (39) ($b_0<0$) and check against (12) by multiplying out the converted factors.

If it is difficult to determine the intersection of the index-line with the β-curve accurately enough (when c_1 and c_2 are small and $b_0>0$) take $\beta = 1$ as the first approximation if $0\leqslant c_2<2$, $\beta = -1$ as the first approximation if $0\geqslant c_2>-2$ and use (26) to obtain the first approximation to β otherwise. Then use (20) to obtain α, and (19) and (20) alternately to obtain improved values of β and α.

To improve the accuracy, start with the approximation obtained graphically, and improve it by one or more applications of either of the iterative processes discussed in Section 4.

If b_0 is very small, say of order 10^{-4}, $x\approx -a_3/4$ is one root, and the remaining roots can be found from (14) by neglecting b_0, dividing through by y, and solving the resulting cubic as indicated in Section 2 - it will be ready-normalized in the form (3).

6. REFERENCES

1. d'OCAGNE, M.: Traite de Nomographie, 1921.

2. d'OCAGNE, M.: Sur la résolution nomographique de l'équation du septième degré. *C.R. Acad. Sci., Paris*, 1900, **131**, pp. 522 - 524.

3. d'OCAGNE, M.: Sur la résolution nomographique des équations algébriques. *Nouv. Annls. Math.*, 1902, 4th series, **2**, pp. 49 - 57.

4. R.R.D.E. Research Report No. 274, 1945. See also:
PORTER, A. and MACK, C.: 'New Methods for the Numerical Solution of Algebraic Equations', *Phil. Mag.*, 1949, 7th series, **40**, pp. 578 - 585, and MACK, C.: 'Routh Test Function Methods for the Numerical Solution of Polynomial Equations', *Q.Jl.Mech. Appl. Math.*, 1959, **12**, pp. 365 - 378.

5. *(Bairstow's Process)* HILDEBRAND, F.B.: 'Introduction to Numerical Analysis', (McGraw-Hill, New York, 1956), p.472.

6. *(Lin's Process)* HEAD, J.W.: Mathematical Techniques in Electronics and Engineering Analysis, (Iliffe, 1964) Chapter 4.

7. GIET, A.: 'Abacs or Nomograms', English edition (translated by J.W. Head and H.D. Phippen) (Iliffe, 1956).

SMW

Printed by BBC Research Department, Kingswood Warren, Tadworth, Surrey